潜变量交互效应标准化估计：方法比较与选用策略*

温忠麟 欧阳劲樱 方俊燕

(华南师范大学心理学院/心理应用研究中心, 广州 510631)

摘 要 标准化估计对模型的解释和效应大小的比较有重要作用。虽然潜变量交互效应的恰当标准化估计公式已经面世超过 10 年，国内外都在使用和引用，但至今未见到关于不同估计方法得到的恰当标准化估计的系统比较。通过模拟实验，比较了乘积指标法、潜调节结构方程(LMS)、无先验信息和有先验信息的贝叶斯法的潜变量交互效应标准化估计在不同条件下的表现。结果发现，在正态条件下，LMS 和有信息贝叶斯法表现较好；而在非正态条件下，乘积指标法比较稳健，但需要较大的样本(不小于 500)，小样本且外生潜变量之间相关较低时可使用无信息贝叶斯法。

关键词 潜变量, 交互效应, 乘积指标法, 潜调节结构方程, 贝叶斯法, 标准化估计

分类号 B841

1 引言

在心理学和其他社会科学领域，结构方程模型可以分析多个潜变量之间的联系。除了线性关系外，交互作用可以揭示变量之间更复杂的关系。因此，分析潜变量的交互效应就成了理论研究和实证研究的一个重要课题。

由于原始数据估计的参数(称为原始估计)会受到变量测量单位的影响，因此不能直接用来比较。研究者常常将参数的原始估计转换为尺度不变的标准化估计。在结构方程模型中，完全标准化估计将指标和潜变量都标准化，使得估计出来的参数(通常感兴趣的是路径系数或因子负荷)可以进行比较。对于潜变量交互效应模型，标准化估计将模型中的潜变量都视为 Z 分数，不仅方便解释交互效应，还可以方便比较简单主效应。

对于显变量情形，可以使用 SPSS 进行交互效应分析，但其输出的标准化估计是不恰当的，而应当使用标准化变量(即 Z 分数)去建模估计(温忠麟等, 2008)。对于包含交互效应的潜变量模型，标准化估计更加复杂。过去很长一段时间，潜变量建模软件(如 Mplus)得到的交互效应标准化估计输出结果是不恰当的，温忠麟等(2008)给出了恰当标准化估计(appropriate standardized

收稿日期: 2021-08-09

* 国家自然科学基金项目(31771245)资助。

通讯作者: 温忠麟, E-mail: wenzl@scnu.edu.cn

estimation)，并证明了这种标准化估计具有尺度不变性(即不会随测量单位变化，Wen et al., 2010)，且在有均值结构或无均值结构的模型中都适用(吴艳等, 2011)。

对于潜变量交互效应标准化估计，目前统计软件输出结果有的是恰当的，有的是不恰当的，需要区别。在不致引起混淆的场合，可以简单地用标准化估计指代恰当标准化估计。Mplus 8.2 (Muthén & Muthén, 2019)及以上版本根据 Wen 等(2010)的公式为潜调节结构方程(Latent Moderated Structural Equations, LMS)和贝叶斯法提供了恰当标准化估计(Asparouhov & Muthén, 2020；也见 Mplus 第 23 号技术文档[①])，使得用户可以非常方便地得到恰当标准化估计。

比较并选择估计方法是潜变量交互效应理论研究的重要组成部分。为解决交互效应项引起的非正态以及非线性等问题，研究者提出了多种建模和估计方法，包括乘积指标法(product indicator approaches, PI)、分布分析方法(distribution analytic approaches，LMS 是其中的一种)、贝叶斯法(Bayesian approaches)等，并在不同情形比较了各种方法的估计偏差(Jackman et al., 2011; Kelava et al., 2011; Marsh et al., 2004; Marsh et al., 2012; 温忠麟, 吴艳, 2010; 温忠麟等, 2013)。但已有的研究未对不同估计方法得到的潜变量交互效应标准化估计进行系统比较，由于 Mplus 8.2 及以上版本提供了恰当标准化估计，为这种比较带来了可行性。注意到恰当标准化估计可以使用原始估计计算出来(Wen et al., 2010; 温忠麟等, 2008)，但原始估计及其相应的标准误都不是尺度不变的(吴艳等, 2014)，有关原始估计的比较结果未必能推论到恰当标准化估计。特别地，恰当标准化估计计算公式中原始估计有的做分子、有的做分母，如果分子分母偏差方向相同且幅度相近，那么恰当标准化估计的偏差有可能很小。因此，尽管已经有文献比较了不同估计方法在潜变量交互效应原始估计上的表现，还是有必要比较不同方法在标准化估计上的表现。

本文假设所论的潜变量及其指标都可以合理地当作连续变量。首先回顾潜变量交互效应的标准化模型、恰当标准化估计及其性质，并梳理潜变量交互效应的主要估计方法以及对它们的比较研究。接着通过一个模拟实验比较乘积指标法、LMS 和贝叶斯法的潜变量交互效应恰当标准化估计在各种常见条件下的表现情况，并给出估计方法选用的建议。

2 潜变量交互效应模型的标准化估计

设潜变量交互效应的结构方程模型为(Marsh et al., 2004)：

$$\eta = \gamma_1\xi_1 + \gamma_2\xi_2 + \gamma_3\xi_1\xi_2 + \zeta \quad (1)$$

其中，ξ_1和ξ_2为外生潜变量，$\xi_1\xi_2$是它们的交互项，η是内生潜变量，ζ是残差项，γ_1，γ_2和γ_3是路径系数(如图 1 所示)。由于方程的截距项及ξ_1和ξ_2的均值不影响交互效应的估计，可以设

[①] http://www.statmodel.com/examples/webnotes/webnote%2023.pdf

定为零(Algina & Moulder, 2001; Marsh et al., 2004)。ξ_1，ξ_2和ζ的均值为零，而$\xi_1\xi_2$的均值通常不为零，故η的均值也通常不为零。

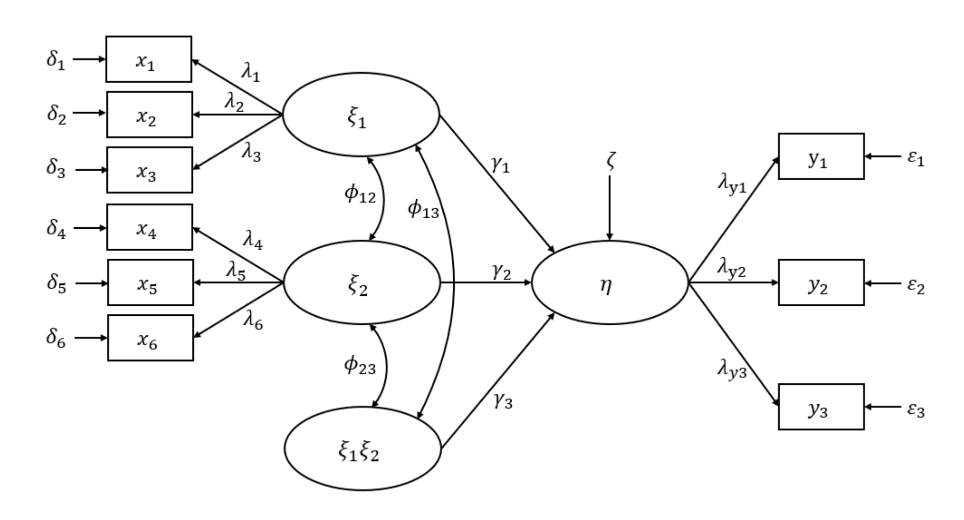


图 1 包含一个交互项的潜变量交互效应模型图

如果按通常的设定，将$\xi_1\xi_2$作为一个变量(而不是两个变量的乘积)去做标准化，方程(1)的标准化形式是

$$\eta' = \gamma_1'\xi_1' + \gamma_2'\xi_2' + \gamma_3'(\xi_1\xi_2)' + \zeta' \tag{2}$$

其中η'是η的标准化变量，ξ_1'和ξ_2'分别是ξ_1和ξ_2的标准化变量，$(\xi_1\xi_2)'$是$\xi_1\xi_2$的标准化变量，ζ'是残差项，γ_1'，γ_2'和γ_3'是通常标准化系数。路径系数的通常标准化估计与原始估计的关系为(Wen et al., 2010; 温忠麟等, 2008)：

$$\gamma_1' = \gamma_1\frac{sd(\xi_1)}{sd(\eta)}, \gamma_2' = \gamma_2\frac{sd(\xi_2)}{sd(\eta)}, \gamma_3' = \gamma_3\frac{sd(\xi_1\xi_2)}{sd(\eta)} \tag{3}$$

其中$sd(\eta)$表示η的标准差的原始估计，余类推。

但方程(2)并不是我们想要的标准化形式，因为它将$\xi_1\xi_2$作为一个变量去做标准化，因而没有了交互乘积项，所得结果难以做交互作用解释。方程(1)的恰当标准化形式为：

$$\eta' = \alpha + \gamma_1''\xi_1' + \gamma_2''\xi_2' + \gamma_3''\xi_1'\xi_2' + \zeta'' \tag{4}$$

其中α是截距项(因为η'是标准化变量，而$\xi_1'\xi_2'$的均值通常不是零，所以截距项一般也不为零)，$\xi_1'\xi_2'$是ξ_1'和ξ_2'交互项，γ_1''，γ_2''和γ_3''是恰当标准化系数。温忠麟等(2008，也见 Wen et al., 2010)推出路径系数的恰当标准化估计与通常标准化估计的关系如下：

$$\gamma_1'' = \gamma_1', \gamma_2'' = \gamma_2',\ \gamma_3'' = \gamma_3'\frac{sd(\xi_1)sd(\xi_2)}{sd(\xi_1\xi_2)} \tag{5}$$

其中用到ξ_1、ξ_2和$\xi_1\xi_2$的标准差的原始估计。如果直接使用原始估计计算恰当标准化估计(Mplus 程序见附录 1)，公式为：

$$\gamma_1'' = \gamma_1\frac{sd(\xi_1)}{sd(\eta)},\ \gamma_2'' = \gamma_2\frac{sd(\xi_2)}{sd(\eta)}, \gamma_3'' = \gamma_3\frac{sd(\xi_1)sd(\xi_2)}{sd(\eta)} \tag{6}$$

上述恰当标准化估计对于潜变量和显变量都适用。由于其推导过程没有涉及具体的估计方法，因此适用于不同的交互效应估计方法，如乘积指标法、分布分析方法、贝叶斯法等。此外，恰当标准化估计与变量的分布无关，无需假设变量是正态分布。

更重要的是，恰当标准化估计具有尺度不变性(Wen et al., 2010)。尺度不变性有三层含义：首先，主效应和交互效应的标准化估计是尺度不变的；第二，所有测量方程中的负荷的标准化估计是尺度不变的；第三，标准化估计相应的标准误和 t 值也是尺度不变的。由于这些性质，使用不同单位的数据，最后得到的标准化估计是相同的。

上述恰当标准化估计在有关潜变量交互效应的研究中已经有许多应用。不少理论研究进行蒙特卡洛模拟时采用了恰当标准化估计(例如，Asparouhov & Muthén, 2020; Büchner & Klein, 2020; Foldnes & Hagtvet, 2014; Kelava et al., 2011; Wu et al., 2013)。Brandt 等(2015)将恰当标准化估计推广到混合模型框架。在人格心理学、发展心理学、临床心理学等领域，不少应用研究也会引用 Wen 等(2010)论文并报告标准化估计结果(例如，Ito et al., 2015; Laczniak et al., 2017; Nagengast et al., 2011; You et al., 2016)。目前流行的结构方程软件 Mplus 在 8.2 或以上版本已经支持 LMS 和贝叶斯法分析潜变量交互效应模型，并提供了恰当标准化估计(Asparouhov & Muthén, 2020)。

3 常用的潜变量交互效应的估计方法

3.1 乘积指标法

乘积指标法最早由 Kenny 和 Judd(1984)提出，使用指标的乘积作为潜变量交互项的指标，包括约束模型、部分约束模型和无约束模型。约束模型和部分约束模型复杂且易出错，只有 Marsh 等(2004)提出的无约束模型被应用工作者接受(温忠麟, 刘红云, 2020)。**故下文如果没有特别说明，乘积指标法使用无约束模型。**

在估计乘积指标模型时，一般选用极大似然(ML; Maximum Likelihood)估计。虽然极大似然估计要求潜变量的指标服从多元正态分布，但在非正态情形依然有较好的稳健性(Boomsma, 1983; Hau & Marsh, 2004)，且可以对标准误和χ^2进行矫正(Satorra & Bentler, 1994)，包括 Mplus 在内的常用 SEM 软件都可以实现(见附录 1)。

3.2 潜调节结构方程

分布分析法可以直接处理交互项$\xi_1\xi_2$的非正态性问题(温忠麟等，2013)，包括潜调节结构方程 (LMS, Klein & Moosbrugger, 2000; Schermelleh-Engel et al., 1998)和准极大似然法(Quasi-Maximum Likelihood, QML, Klein & Muthén, 2007)。

目前只有 LMS 可以在 Mplus 实现，而 QML 方法需要使用专门的软件，故本文只考虑 LMS。LMS 将指标的联合分布近似为有限个条件正态分布的混合分布，并通过 EM 算法迭代得到参数的极大似然估计值，是一种广义的极大似然方法。Klein 和 Moosbrugger(2000)的研究表明 LMS 对潜变量交互效应的估计非常有效。但值得注意的是，LMS 虽然注意到了交互项$\xi_1\xi_2$的非正态性，但依然建立在潜变量正态分布的假设上，在潜变量ξ_1和ξ_2非正态分布时会产生较大的偏差(Cham et al.,2012; Kelava & Nagengast, 2012)。

3.3 贝叶斯法

在结构方程模型中引入贝叶斯法也可以用来估计潜变量交互效应(Arminger & Muthén, 1998; Lee et al., 2007)。贝叶斯结构方程模型结合先验分布与给定数据的似然函数确定未知参数的后验分布，并使用马尔科夫链蒙特卡洛方法，如 Gibbs 采样(Geman & Geman, 1984)和 Metropolis–Hastings(MH)算法(Hastings, 1970; Metropolis et al., 1953)估计后验分布。贝叶斯法不仅对简单的结构方程模型有可信的估计，对非线性模型的估计同样可以接受(Lee & Song, 2004; Lee et al., 2007)。过去潜变量交互效应模型的贝叶斯法通常使用 WinBugs 软件实现，而 Mplus 8.2 及以上版本已经可以实现，还可以给出恰当标准化估计。

贝叶斯法常被认为在小样本、复杂模型如变量非线性关系等情况下表现出优势(张沥今等, 2019)。由于贝叶斯法结合了先验信息与观测数据，随着样本容量减小，先验信息对后验分布的影响增强而数据对其影响减弱，因而贝叶斯法无需大样本渐近理论(asymptotic theory)，在小样本条件有天然的优势(Gelman et al., 2014; Lee & Song, 2004; Lee et al., 2007; Muthén & Asparouhov, 2012)。Lee 等(2007)的模拟研究发现贝叶斯法在小样本(N = 150)下对潜变量交互效应的估计也可接受。但贝叶斯法的小样本优势并非是绝对的：在无先验信息时，贝叶斯法可能会出现严重的偏差，比极大似然估计更差，只有在先验信息准确时，贝叶斯法的表现才可能在小样本下可接受(Smid et al., 2020)。而在样本容量很大时，贝叶斯估计与极大似然估计等价(Gelman et al., 2014)。Lee 和 Song(2004)在小样本时估计验证性因子分析和结构方程模型发现，贝叶斯法在非正态情形并不稳健。Jia(2016)的研究发现，贝叶斯结构方程在严重非正态情形下参数置信区间的真值覆盖率过低。

3.4 潜变量交互效应的估计方法的比较

对于原始估计，已有文献比较了乘积指标法、LMS 和贝叶斯法估计潜变量交互效应的表现。在正态或偏态很小的情形，无约束模型在样本容量较大时估计偏差、标准误偏差和统计检验力可以接受；而 LMS 的表现相对更好，在小样本时估计偏差和标准误偏差也可以忽略，且统计检验力更高 (Cham et al., 2012; Jackman et al., 2011; Marsh et al., 2004)。而在非正态情形，无约束模型在大样本时的估计偏差、标准误偏差、覆盖率和 I 类错误率都可接受，而 LMS 对非线

性效应及其标准误的估计有偏，且比无约束模型有更高的I类错误率(Cham et al., 2012; Kelava & Nagengast, 2012; Marsh et al., 2004)。

目前对贝叶斯法与其他两种方法进行系统比较的研究很少。 Lee 等(2004)比较了贝叶斯法与有约束模型的乘积指标法，发现在潜变量正态分布下，贝叶斯法在小样本下也表现良好，而有约束模型在样本较大时才能接受。Kelava 和 Nagengast(2012)系统比较了混合模型框架下的贝叶斯法与无约束模型的乘积指标法和 LMS 在潜变量不同分布下的表现，发现在潜变量非正态时使用二到四个潜类别的贝叶斯法对非线性效应的估计无偏，优于其他方法。Asparouhov 和 Muthén(2020)在多水平框架下比较了 LMS 和无先验信息的贝叶斯法在正态大样本(N = 1000)的表现，结果表明贝叶斯法较好。

总之，在正态情形 LMS 比无约束乘积指标法更好；在非正态情形，LMS 有较大的偏差，无约束的乘积指标法在大样本时估计偏差可以接受；温忠麟和刘红云(2020)推测贝叶斯法可能在小样本和非正态条件下有交好的表现，但需要进一步的研究。

但由于潜变量交互效应恰当标准化估计是根据原始估计(或者连同通常标准化估计)使用公式计算出来的，而不是某个模型的直接估计结果，所以上述有关估计方法在原始估计的表现与在恰当标准化估计的表现未必一致。**无论一致与否，都需要专门研究才能得出结论；无论一致与否，都是很有意义的结果。**下面将系统比较潜变量交互效应模型的乘积指标法、LMS、有先验信息的贝叶斯法(Bayesian approaches with informative priors, BI)和无先验信息的贝叶斯法(Bayesian approaches with noninformative priors, BN)在各种常见条件下的恰当标准化估计的表现，为应用研究选择估计方法提供参考。

4 潜变量交互效应标准化估计：乘积指标法、LMS 和贝叶斯法的比较研究

4.1 模拟研究设计

本研究考察外生潜变量在不同的分布、相关大小、样本容量条件下，乘积指标法、LMS和贝叶斯法得到的潜变量交互效应恰当标准化估计的表现。参考 Marsh 等(2004)的模拟研究设计，产生数据的结构模型(真模型)为

$$\eta = \gamma_1\xi_1 + \gamma_2\xi_2 + \gamma_3\xi_1\xi_2 + \zeta$$

其中，内生潜变量η的方差设定为 1，两个外生潜变量ξ_1，ξ_2是标准化变量(M = 0, SD = 1)，因而真模型的主效应和交互效应都是标准化的(不受η均值的影响)。

考虑两种产生ξ_1，ξ_2的分布条件：正态条件和非正态条件，其中非正态分布由两个$\chi^2(6)$分布加权得到 (Marsh et al., 2004)。两种条件下潜变量的偏度和峰度见表 1 和表 2。潜变量η，

ξ_1，ξ_2 各有 3 个指标，分别是y_1，y_2，y_3；x_1，x_2，x_3和x_4，x_5，x_6。指标在因子上的标准化负荷都是 0.7，固定负荷的合理性在于前人的研究发现负荷大小基本上不会影响估计方法的比较结果(例如，Marsh et al., 2004)。所有的指标都是由因子ξ_1，ξ_2加上与它们分布条件相同的随机误差得到。

模型中的两个主效应γ_1和γ_2都固定为 0.4，交互效应γ_3为 0(计算I类错误率)或 0.2(计算检验力)。外生潜变量ξ_1和ξ_2的相关设为 0，0.3 或 0.7。交互效应γ_3为 0.2 时，对于正态分布，不同潜变量相关大小(0，0.3，0.7)对应的交互效应项的效应量ρ^2 (Marsh et al., 2004；即交互项的 R^2 增加量)分别是4.00%，4.36%，5.96%。对于非正态分布，ρ^2分别是4.00%，5.07%，9.88%。常见的交互效应项的效应量一般在 3%到 8%之间(Champoux & Peters, 1987; Moulder & Algina, 2002)，说明研究设计是合理的。

样本容量设定 $N = 100$，200，500，这是应用研究典型的样本容量(Jaccard & Wan, 1995; Lee et al., 2007; MacCallum & Austin, 2000; Marsh et al., 2004)。

综上所述，模拟研究的因素设计如下：

ξ_1和ξ_2的分布：正态分布，非正态分布；(被试间因素)

ξ_1和ξ_2的相关：$\phi_{12} = 0$，0.3，0.7；(被试间因素)

交互效应：$\gamma_3 = 0$，0.2；(被试间因素)

样本容量：$N = 100$，200，500；(被试间因素)

估计方法：乘积指标法，LMS，无先验信息的贝叶斯法，有先验信息的贝叶斯法。(被试内因素)

这是一个 2×3×2×3×4 设计，共有 144 种处理，被试间处理共 36 种。对于每种被试间处理，使用 PRELIS 2.3(Jöreskog & Sörbom, 1996)产生 500 个样本(即重复 500 次)。模拟产生的初始指标方差都是 1，因子负荷按前述设计，测量误差的分布与产生ξ_1，ξ_2的分布条件相同。参考 Marsh 等(2004)做法，对模拟得到的数据再作下面的线性变换，以变换成不同的测量单位：

$$y_1' = 1 + y_1/0.7,\ y_2' = 2 + y_2,\ y_3' = 3 + y_3;$$
$$x_1' = 7 + x_1/0.7,\ x_2' = 6 + x_2,\ x_3' = 8 + 2x_3;$$
$$x_4' = 6 + x_4/0.7,\ x_5' = 9 + 3x_5,\ x_6' = 7 + x_6;$$

由此变换后的指标得到包含截距项的原始数据。在建立潜变量交互效应模型时，首先将所有指标进行中心化。对每一个样本，使用 Mplus 8.2 对潜变量交互效应模型进行估计，分别得到乘积指标法、LMS、无先验信息和有先验信息的贝叶斯法估计的结果。

乘积指标法使用无均值结构的模型(Lin et al., 2010; Marsh et al., 2004; 吴艳等, 2009)，采用配对乘积指标x_1x_4，x_2x_5，x_3x_6(Marsh et al., 2004)，并使用固定负荷法指定测量单位。无论潜变量是何分布，乘积指标都不是正态分布，在 Mplus 中指定“ESTIMATOR=MLM”实现对参数

估计的标准误和模型卡方统计量的矫正(Satorra & Bentler, 1994)。根据温忠麟等(2008，也见 Wen et al., 2010)的公式利用 Mplus 中“Model Constraint”命令将原始估计转化为恰当标准化估计(Mplus 程序见附录 1)。

LMS 的恰当标准化估计已经内置于 Mplus 8.2 及以上版本，只需要输出标准化解便可，不需要重新计算。Mplus 程序参见温忠麟和刘红云(2020, P193-194)。

贝叶斯法采用 Mplus 的默认设置 (Asparouhov & Muthén, 2020)。在有先验信息的条件下，设定因子负荷、主效应和交互效应的先验分布服从均值为真值，方差为 0.01 的正态分布(Muthén & Asparouhov, 2012)。潜变量ξ_1，ξ_2的方差协方差矩阵的先验分布设定为逆 Wishart 分布，其均值为真值，对角线元素方差设定为 0.01，由于逆 Wishart 分布中一个元素的信息会决定其他元素的信息(Gelman et al., 2014)，非对角线元素方差在 0.50 到 0.74 之间，计算公式见 Asparouhov 和 Muthén(2010)。贝叶斯法的恰当标准化估计已经内置于 Mplus 8.2 及以上版本，只需要输出标准化解便可，不需要重新计算。Mplus 程序参见温忠麟和刘红云(2020, P199-200)。

下面报告的模拟结果所说的标准化估计指的是恰当标准化估计，交互效应真值$\gamma_3 = 0$ 的结果见表 1 和表 2，$\gamma_3 = 0.2$ 的结果见表 3 和表 4。

4.2 模拟研究结果

4.2.1 适当解

适当解是指模型收敛且所有参数估计都合理的解(例如，方差和标准误为非负值)。参数估计偏差、标准误偏差等其他指标的计算只基于适当解的结果。从表 1~表 4 可以看出，LMS 和两种贝叶斯法在所有情况下适当解的比例都接近或等于 100％。对于乘积指标法，$N = 100$ 时适当解的比例从 $\phi_{12} = 0$(平均 69.6%)、$\phi_{12} = 0.3$(平均 79.1%)到$\phi_{12} = 0.7$(平均 88.5%)逐渐提高；在 $N = 200$ 时适当解的比例超过 90%(潜变量非正态且$\phi_{12} = 0$ 时除外)；在 $N = 500$ 时全部接近或等于 100％。

4.2.2 交互效应标准化估计的偏差

当$\gamma_3 = 0$ 时，交互效应标准化估计的偏差等于交互效应标准化估计的均值。偏差不超过 0.02 认为可以接受（后面$\gamma_3 = 0.2$ 情形，相对偏差 10%对应的偏差即为 0.02）。潜变量ξ_1，ξ_2正态分布时(表 1)， 乘积指标法和 LMS 的偏差(正文中谈到偏差均指绝对值)均小于 0.01；两种贝叶斯法的偏差均小于 0.02。总之，当交互效应不存在时，在正态条件下，无论哪种方法，其标准化估计都很接近零。

潜变量ξ_1，ξ_2非正态分布时(表 2)，乘积指标法在$\phi_{12} = 0$ 且 $N = 100$ 时交互效应的估计偏差为 0.031，其他情形均小于 0.01；LMS 和两种贝叶斯法在非正态情形没有乘积指标法那么稳健，交互效应标准化估计偏差随着ϕ_{12}增大(因而$\xi_1\xi_2$的偏度和峰度也增大)而增大，并且不随样

本容量的增大而系统减少，除了$\phi_{12}=0$外，偏差都超过 0.02。从方法比较的角度来说，在非正态条件下，除了 $\phi_{12}=0$ 且 $N=100$ 外，乘积指标法偏差最小，排序大多为 PI < BI < LMS < BN。

表 1 正态分布下潜变量交互效应的标准化估计结果($\gamma_3=0$)

N	方法	适当解%	M	SD	SE	SE 偏差%	第 I 类错误率
			$\phi_{12}=0$, ξ_1 (0.001, −0.004), ξ_2 (−0.008, −0.007), $\xi_1\xi_2$ (−0.024, 6.163)				
100	PI	72.8	−0.009	0.239	0.180	−24.5	0.069
	LMS	100.0	−0.003	0.148	0.147	−0.8	0.082
	BN	100.0	−0.005	0.146	0.147	0.6	0.060
	BI	100.0	−0.003	0.047	0.083	75.8	0.000
200	PI	92.2	0.001	0.137	0.116	−15.3	0.043
	LMS	100.0	0.005	0.105	0.100	−4.8	0.080
	BN	100.0	0.019	0.107	0.098	−8.5	0.088
	BI	100.0	0.011	0.053	0.068	28.3	0.010
500	PI	100.0	0.000	0.077	0.072	−6.6	0.052
	LMS	100.0	0.002	0.064	0.063	−1.6	0.060
	BN	100.0	0.001	0.065	0.063	−3.1	0.058
	BI	100.0	0.002	0.046	0.053	14.1	0.030
			$\phi_{12}=0.3$, ξ_1 (−0.001, 0.000), ξ_2 (−0.001, 0.001), $\xi_1\xi_2$ (1.619, 7.628)				
100	PI	77.4	0.006	0.207	0.163	−21.4	0.059
	LMS	100.0	0.006	0.136	0.129	−4.8	0.082
	BN	100.0	0.008	0.133	0.135	1.1	0.044
	BI	100.0	0.003	0.049	0.081	65.3	0.000
200	PI	92.4	0.009	0.121	0.105	−13.1	0.045
	LMS	100.0	0.003	0.092	0.089	−3.4	0.070
	BN	100.0	0.016	0.093	0.088	−5.7	0.068
	BI	100.0	0.010	0.051	0.064	24.8	0.008
500	PI	99.6	0.004	0.064	0.063	−0.4	0.034
	LMS	100.0	0.001	0.054	0.055	1.2	0.048
	BN	100.0	0.001	0.056	0.055	−1.5	0.054
	BI	100.0	0.001	0.042	0.048	14.4	0.020
			$\phi_{12}=0.7$, ξ_1 (−0.005, 0.002), ξ_2 (−0.005, −0.002), $\xi_1\xi_2$ (2.673, 11.073)				
100	PI	84.2	0.009	0.153	0.120	−21.9	0.076
	LMS	99.8	0.008	0.099	0.094	−4.5	0.078
	BN	96.2	0.011	0.100	0.103	3.5	0.037
	BI	100.0	0.006	0.048	0.072	49.1	0.006
200	PI	97.2	0.008	0.089	0.079	−10.6	0.045
	LMS	100.0	0.003	0.065	0.065	0.2	0.054
	BN	100.0	0.013	0.067	0.065	−2.1	0.062
	BI	100.0	0.012	0.045	0.054	20.5	0.034
500	PI	100.0	0.004	0.049	0.048	−2.3	0.046
	LMS	100.0	0.002	0.040	0.041	0.9	0.048
	BN	100.0	0.001	0.041	0.040	−2.3	0.050
	BI	100.0	0.000	0.035	0.037	7.0	0.032

注：适当解 %= 收敛到适当解的比率(其他结果都基于适当解)，M = 交互效应标准化估计的均值，SD = 交互效应标准化估计的标准差，SE = 交互效应的标准误的均值。PI = 乘积指标法，LMS = 潜调节结构方程，BN = 无先验信息的贝叶斯法，BI = 有先验信息的贝叶斯法。$\phi_{12}=\xi_1$和ξ_2的协方差。ξ_1，ξ_2和$\xi_1\xi_2$括号中的数值为分布的偏度和峰度，使用模拟生成的全部 250000 = 500 × 500 个数据计算得到。

表 2 非正态分布下潜变量交互效应的标准化估计结果(γ_3 = 0)

N	方法	适当解%	M	SD	SE	SE 偏差%	第 I 类错误率
			ϕ_{12} = 0, ξ_1 (1.141, 1.962), ξ_2 (1.152, 2.002), $\xi_1\xi_2$ (1.278, 20.126)				
100	PI	64.6	0.031	0.239	0.198	−17.0	0.043
	LMS	99.4	0.015	0.155	0.158	2.1	0.078
	BN	100.0	0.020	0.153	0.153	−0.3	0.064
	BI	100.0	0.005	0.047	0.083	77.1	0.004
200	PI	83.2	0.009	0.157	0.130	−17.2	0.060
	LMS	100.0	0.013	0.112	0.108	−3.9	0.078
	BN	100.0	0.030	0.114	0.102	−10.1	0.096
	BI	100.0	0.017	0.054	0.068	26.5	0.034
500	PI	97.8	0.001	0.087	0.077	−11.5	0.059
	LMS	100.0	0.016	0.067	0.067	−0.9	0.092
	BN	100.0	0.017	0.069	0.064	−7.3	0.088
	BI	100.0	0.013	0.047	0.053	12.9	0.034
			ϕ_{12} = 0.3, ξ_1 (0.855, 1.121), ξ_2 (0.858, 1.161), $\xi_1\xi_2$ (3.200, 32.430)				
100	PI	75.8	0.008	0.194	0.155	−20.0	0.069
	LMS	100.0	0.049	0.135	0.132	−2.1	0.088
	BN	100.0	0.058	0.135	0.137	1.6	0.060
	BI	100.0	0.021	0.049	0.081	64.7	0.010
200	PI	91.6	0.005	0.113	0.104	−8.2	0.057
	LMS	100.0	0.044	0.090	0.090	−0.2	0.068
	BN	100.0	0.061	0.092	0.088	−4.3	0.122
	BI	100.0	0.035	0.051	0.063	25.4	0.046
500	PI	100.0	−0.002	0.069	0.064	−7.8	0.056
	LMS	100.0	0.044	0.056	0.055	−1.3	0.122
	BN	100.0	0.046	0.057	0.055	−4.5	0.136
	BI	100.0	0.034	0.043	0.047	11.4	0.088
			ϕ_{12} = 0.7, ξ_1 (0.850, 1.075), ξ_2 (0.856, 1.110), $\xi_1\xi_2$ (6.275, 86.323)				
100	PI	95.0	0.002	0.137	0.118	−14.0	0.072
	LMS	100.0	0.064	0.100	0.098	−2.4	0.106
	BN	97.2	0.074	0.101	0.107	5.8	0.080
	BI	100.0	0.035	0.048	0.071	48.2	0.032
200	PI	100.0	0.000	0.081	0.074	−8.8	0.058
	LMS	100.0	0.059	0.067	0.064	−3.8	0.136
	BN	99.6	0.076	0.069	0.067	−3.4	0.177
	BI	100.0	0.053	0.045	0.053	17.0	0.118
500	PI	100.0	−0.001	0.046	0.043	−6.4	0.066
	LMS	100.0	0.056	0.041	0.040	−2.5	0.272
	BN	100.0	0.057	0.042	0.040	−5.8	0.292
	BI	100.0	0.048	0.035	0.036	2.7	0.248

注：适当解 %= 收敛到适当解的比率(其他结果都基于适当解)，M = 交互效应标准化估计的均值，SD = 交互效应标准化估计的标准差，SE = 交互效应的标准误的均值。PI = 乘积指标法，LMS = 潜调节结构方程，BN = 无先验信息的贝叶斯法，BI = 有先验信息的贝叶斯法。ϕ_{12} = ξ_1和ξ_2的协方差。ξ_1，ξ_2和$\xi_1\xi_2$括号中的数值为分布的偏度和峰度，使用模拟生成的全部 250000 = 500 × 500 个数据计算得到。

这里重点考察γ_3 = 0.2 的情形(表 3 和表 4)。交互效应标准化估计的相对偏差等于 500 个样本标准化估计的均值减去真值后的值再除以真值。一般认为，相对偏差(绝对值)小于 5%时可以

忽略(Hoogland & Boomsma, 1998)，相对偏差不超过 10%(Cham et al., 2012)或 15%(Bandalos, 2002; Muthén et al., 1987)时可以接受。

潜变量ξ_1，ξ_2正态分布时(表 3)，对于乘积指标法，在 N = 100 且(ϕ_{12} = 0.3 或 0.7)、以及 N = 200 且ϕ_{12} = 0.3 的情形相对偏差超过 10%，在其他情形均小于 10%。LMS 的相对偏差在所有情形均接近或小于 5%。无先验信息的贝叶斯法的相对偏差在所有情形均小于 10%。有先验信息的贝叶斯法的相对偏差在所有情况下均接近或小于 5%。从方法之间比较的角度来说，在 N = 100 时，有先验信息的贝叶斯法相对偏差最小，乘积指标法的相对偏差最大，排序为 BI < LMS, BN ≤ PI；在 N = 200 或 500 时，LMS 的相对偏差最小，排序大多为 LMS < BI < BN < PI。总之，在正态情形，LMS 和 BI 表现比较好，其次是 BN，而乘积指标法需要较大的样本(如 N = 500)才可接受。

潜变量ξ_1，ξ_2非正态分布时(表 4)，对于乘积指标法，N = 100 且(ϕ_{12} = 0 或 0.3)时，交互效应的相对偏差大于 10%，在其他情形接近或小于 10%。LMS、两种贝叶斯法的稳健性远远不如乘积指标法，交互效应标准化估计相对偏差随着ϕ_{12}增大(因而$\xi_1\xi_2$的偏度和峰度也增大)而增大，并且不随样本容量的增大而系统减少：在ϕ_{12} = 0 时相对偏差均接近或小于 10%，而ϕ_{12} = 0.3 时均大于 10%(甚至超过 30%)，ϕ_{12} = 0.7 时均大于 30%(甚至超过 50%)。总之，在潜变量ξ_1，ξ_2非正态分布时，ξ_1和ξ_2的相关越大，非正态程度越严重，只要ξ_1和ξ_2有实质性相关，就应当避免使用 LMS 和贝叶斯法，此时只有乘积指标法可以使用，但要有较大的样本容量(如 N = 500)。

4.2.3 交互效应标准化估计的标准误的估计偏差

每个实验处理中，将重复抽取的 500 个样本得到的交互效应标准化估计的标准差(记为 *SD*)作为标准误的真值，交互效应的标准误的相对偏差等于交互效应标准误的均值(记为 *SE*) 减去真值后的值再除以真值，相对偏差(绝对值)在 10%以内可以接受(Hoogland & Boomsma, 1998)。考虑到模型的交互项总是非正态的，对于乘积指标法，标准误使用 Satorra-Bentler 法进行矫正(Satorra & Bentler, 1994)。结果见表 1~表 4。

先看γ_3 = 0 的情形。潜变量ξ_1，ξ_2正态分布时(表 1)，乘积指标法在 N = 100 或 N = 200 时标准误的相对偏差均大于 10%，在 N = 500 时小于 10%；LMS 和无信息贝叶斯法的标准误的相对偏差在所有情形均小于 10%；除了在 N = 500 且ϕ_{12} = 0.7 时，有信息贝叶斯法的标准误的相对偏差均大于 10%。总之，交互效应标准化估计的标准误的相对偏差是 LMS 和 BN 的较小，且都可以接受，乘积指标法的在样本较大时才能接受，而 BI 基本上在各种条件下都是远远高估的。

潜变量ξ_1，ξ_2非正态分布时(表 2)，乘积指标法在小样本或ξ_1和ξ_2低相关时，即 N = 100 或ϕ_{12} = 0 时的 5 种情形，交互效应标准误的相对偏差大于 10%，其他 4 种情形小于 10%；LMS 和无信息贝叶斯法标准误的相对偏差在所有情形均接近或小于 10%；除了在 N = 500 且ϕ_{12} = 0.7

时，有信息贝叶斯法的标准误的相对偏差均大于 10%。总之，在非正态条件下，各种方法的交互效应标准化估计的标准误的相对偏差与正态情形类似，即 LMS 和 BN 的较小且都可以接受，乘积指标法的在样本较大时才能接受，而 BI 基本上在各种条件下都是远远高估的。

表 3 正态分布下潜变量交互效应的标准化估计结果(γ_3 = 0.2)

N	方法	适当解%	M	M 偏差%	SD	SE	SE 偏差%	检验力
ϕ_{12} = 0, ξ_1 (0.001, −0.004), ξ_2 (−0.008, −0.007), $\xi_1\xi_2$ (−0.024, 6.163)								
100	PI	73.8	0.217	8.3	0.216	0.183	−15.6	0.236
	LMS	100.0	0.189	−5.7	0.148	0.144	−2.6	0.300
	BN	100.0	0.183	−8.3	0.146	0.145	−0.2	0.268
	BI	100.0	0.195	−2.6	0.051	0.084	65.0	0.766
200	PI	95.2	0.213	6.7	0.141	0.129	−8.8	0.416
	LMS	100.0	0.199	−0.5	0.104	0.098	−6.3	0.566
	BN	100.0	0.215	7.3	0.106	0.095	−10.4	0.638
	BI	100.0	0.208	4.1	0.054	0.067	24.1	0.924
500	PI	100.0	0.204	2.1	0.081	0.079	−2.4	0.774
	LMS	100.0	0.201	0.3	0.061	0.061	0.0	0.878
	BN	100.0	0.202	1.2	0.062	0.061	−1.7	0.888
	BI	100.0	0.201	0.7	0.045	0.051	13.2	0.986
ϕ_{12} = 0.3, ξ_1 (−0.001, 0.000), ξ_2 (−0.001, 0.001), $\xi_1\xi_2$ (1.619, 7.628)								
100	PI	81.2	0.230	15.2	0.210	0.176	−15.9	0.273
	LMS	100.0	0.198	−1.1	0.133	0.124	−7.0	0.420
	BN	100.0	0.200	−0.2	0.130	0.131	1.1	0.374
	BI	100.0	0.200	−0.2	0.051	0.081	58.5	0.778
200	PI	95.8	0.226	12.9	0.142	0.120	−15.3	0.476
	LMS	100.0	0.201	0.7	0.089	0.086	−3.4	0.626
	BN	100.0	0.215	7.6	0.089	0.085	−4.7	0.674
	BI	100.0	0.209	4.6	0.051	0.063	21.9	0.962
500	PI	99.6	0.207	3.7	0.072	0.071	−1.4	0.871
	LMS	100.0	0.201	0.7	0.054	0.054	0.1	0.954
	BN	100.0	0.204	2.1	0.055	0.053	−3.5	0.964
	BI	100.0	0.202	1.0	0.042	0.046	9.3	0.998
ϕ_{12} = 0.7, ξ_1 (−0.005, 0.002), ξ_2 (−0.005, −0.002), $\xi_1\xi_2$ (2.673, 11.073)								
100	PI	89.6	0.232	15.9	0.165	0.134	−18.7	0.446
	LMS	99.8	0.203	1.4	0.095	0.090	−5.4	0.623
	BN	97.0	0.206	3.2	0.094	0.099	5.4	0.555
	BI	100.0	0.202	0.9	0.048	0.070	46.2	0.906
200	PI	98.2	0.218	8.8	0.104	0.089	−14.0	0.735
	LMS	100.0	0.202	0.9	0.063	0.063	−0.7	0.878
	BN	100.0	0.214	7.1	0.063	0.063	−0.9	0.922
	BI	100.0	0.212	5.8	0.045	0.052	15.8	0.992
500	PI	100.0	0.206	3.0	0.058	0.054	−5.7	0.982
	LMS	100.0	0.201	0.6	0.039	0.039	−0.1	0.998
	BN	100.0	0.203	1.4	0.040	0.038	−5.7	1.000
	BI	100.0	0.202	0.8	0.034	0.035	3.0	1.000

注：适当解 %= 收敛到适当解的比率(其他结果都基于适当解)，M = 交互效应标准化估计的均值，SD = 交互效应标准化估计的标准差，SE = 交互效应的标准误的均值。PI = 乘积指标法，LMS = 潜调节结构方程，BN = 无先验信息的贝叶斯法，BI = 有先验信息的贝叶斯法。ϕ_{12} = ξ_1和ξ_2的协方差。ξ_1，ξ_2和$\xi_1\xi_2$括号中的数值为分布的偏度和峰度，使用模拟生成的全部 250000 = 500 × 500 个数据计算得到。

表 4 非正态分布下潜变量交互效应的标准化估计结果(γ_3 = 0.2)

N	方法	适当解%	*M*	*M* 偏差%	*SD*	*SE*	*SE* 偏差%	检验力
			$\phi_{12} = 0$, ξ_1 (1.141, 1.962), ξ_2 (1.152, 2.002), $\xi_1\xi_2$ (1.278, 20.126)					
100	PI	67.0	0.232	15.8	0.223	0.202	−9.7	0.194
	LMS	99.6	0.197	−1.4	0.152	0.153	0.3	0.329
	BN	100.0	0.201	0.7	0.151	0.148	−2.2	0.314
	BI	100.0	0.200	0.2	0.050	0.084	69.9	0.772
200	PI	88.6	0.223	11.4	0.177	0.149	−15.8	0.348
	LMS	100.0	0.206	2.9	0.106	0.102	−4.4	0.556
	BN	100.0	0.223	11.7	0.107	0.098	−8.5	0.620
	BI	100.0	0.214	6.9	0.055	0.067	23.1	0.920
500	PI	99.0	0.212	6.0	0.101	0.091	−10.3	0.725
	LMS	100.0	0.214	6.8	0.065	0.063	−2.5	0.890
	BN	100.0	0.217	8.7	0.066	0.061	−6.9	0.904
	BI	100.0	0.212	6.1	0.047	0.052	11.0	0.988
			$\phi_{12} = 0.3$, ξ_1 (0.855, 1.121), ξ_2 (0.858, 1.161), $\xi_1\xi_2$ (3.200, 32.430)					
100	PI	81.8	0.239	19.7	0.207	0.173	−16.5	0.330
	LMS	99.8	0.252	25.8	0.131	0.129	−2.0	0.533
	BN	100.0	0.260	30.0	0.129	0.131	1.4	0.528
	BI	100.0	0.227	13.4	0.053	0.082	54.8	0.878
200	PI	95.2	0.223	11.5	0.132	0.117	−11.2	0.527
	LMS	100.0	0.251	25.4	0.089	0.088	−0.5	0.804
	BN	100.0	0.270	34.9	0.089	0.085	−5.1	0.862
	BI	100.0	0.244	22.1	0.054	0.063	17.3	0.986
500	PI	100.0	0.207	3.4	0.075	0.071	−6.1	0.868
	LMS	100.0	0.253	26.7	0.058	0.055	−5.4	0.990
	BN	100.0	0.258	28.9	0.059	0.053	−10.1	0.994
	BI	100.0	0.245	22.7	0.046	0.046	0.6	1.000
			$\phi_{12} = 0.7$, ξ_1 (0.850, 1.075), ξ_2 (0.856, 1.110), $\xi_1\xi_2$ (6.275, 86.323)					
100	PI	85.0	0.221	10.3	0.133	0.114	−13.8	0.576
	LMS	99.8	0.299	49.5	0.098	0.094	−4.6	0.872
	BN	98.2	0.302	51.2	0.094	0.101	7.9	0.851
	BI	100.0	0.267	33.4	0.052	0.073	40.2	0.986
200	PI	96.0	0.210	4.9	0.078	0.073	−6.1	0.831
	LMS	100.0	0.297	48.7	0.065	0.065	−0.2	0.992
	BN	99.4	0.303	51.4	0.064	0.063	−2.6	0.994
	BI	100.0	0.291	45.7	0.047	0.051	8.3	0.998
500	PI	100.0	0.204	1.8	0.047	0.044	−6.1	0.992
	LMS	100.0	0.297	48.6	0.041	0.040	−2.7	1.000
	BN	100.0	0.298	49.2	0.041	0.039	−5.2	1.000
	BI	100.0	0.290	45.1	0.037	0.035	−5.5	1.000

注：适当解 %= 收敛到适当解的比率(其他结果都基于适当解)，*M* = 交互效应标准化估计的均值，*SD* = 交互效应标准化估计的标准差，*SE* = 交互效应的标准误的均值。PI = 乘积指标法，LMS = 潜调节结构方程，BN = 无先验信息的贝叶斯法，BI = 有先验信息的贝叶斯法。$\phi_{12} = \xi_1$和ξ_2的协方差。ξ_1，ξ_2和$\xi_1\xi_2$括号中的数值为分布的偏度和峰度，使用模拟生成的全部 250000 = 500 × 500 个数据计算得到。

再看γ_3 = 0.2 的情形。潜变量ξ_1，ξ_2正态分布时(表 3)，四种方法交互效应标准误的相对偏差与γ_3 = 0 时的情形类似，不再赘述。

潜变量ξ_1，ξ_2非正态分布时(表 4)，乘积指标法在 *N* = 100 且ϕ_{12} = 0 时，*N* = 200 且ϕ_{12} = 0.3(或 0.7)时，以及 *N* = 500 时交互效应标准误的相对偏差小于 10%，在其他情形大于 10%；LMS 和无信息贝叶斯法在所有情形标准误的相对偏差均接近或小于 10%；有信息贝叶斯法在 *N*

= 200 且ϕ_{12} = 0.7 时，以及 N = 500 时交互效应标准误的相对偏差小于 10%，在其他情形接近或大于 10%。总之，在非正态条件下，各种方法的交互效应标准化估计的标准误的相对偏差，LMS 和 BN 的较小且都可以接受，乘积指标法和 BI 的在样本较大时才能接受。

4.2.4 交互效应检验的I类错误率和统计检验力

因为检验有尺度不变性，对标准化估计的检验结果与对原始估计的检验结果应当相同或者相近，所以不用强调是对标准化估计的检验。检验的I类错误率指零假设为真 (这里是γ_3 = 0) 时，错误地拒绝它的概率，其值等于在 0.05 水平上 500 个样本拒绝零假设的比例。I类错误率低于 0.075 是可接受的标准(Bradley, 1978; Wu et al., 2013)，结果见表 1 和表 2。

潜变量ξ_1，ξ_2正态分布时(表 1)，乘积指标法在所有情形的I类错误率均接近或低于 0.075；LMS 在 N = 100 时的 3 种相关条件下，以及 N = 200 且ϕ_{12} = 0 时略高于 0.075，在其他 5 种情形低于 0.075；除了在 N = 200 且ϕ_{12} = 0 时，无信息贝叶斯法的I类错误率均低于 0.075；有信息贝叶斯法在所有情形的I类错误率均低于 0.075。总之，LMS 在多数情况下有较高的 I 类错误率，在样本容量较大时可以接受；其他方法的 I 类错误率可以接受，其中有信息贝叶斯法的最低。

潜变量ξ_1，ξ_2非正态分布时(表 2)，乘积指标法在所有情形的I类错误率均低于 0.075；LMS 在 N = 200 且ϕ_{12} = 0.3 时低于 0.075，其他 8 种情形均高于 0.075；无信息贝叶斯法在 N = 100 且(ϕ_{12} = 0 或 0.3)时低于 0.075，其他 7 种情形均高于 0.075；有信息贝叶斯法在 N = 100 的 3 种相关条件下，N = 200 且ϕ_{12} = 0 或 0.3 时，以及 N = 500 且ϕ_{12} = 0 时低于 0.075，其他 3 种情形高于 0.075。总之，除了乘积指标法外，其他方法的 I 类错误率倾向于在小样本或交互项非正态性不严重(ξ_1和ξ_2相关较小)时才可以接受。

参数的统计检验力指零假设为假(这里是γ_3 = 0.2)时，正确拒绝它的概率，其值等于在 0.05 水平上 500 个样本正确地拒绝零假设的比例，见表 3 和表 4。统计检验力越接近 1 越好，通常希望高于 80%。

潜变量ξ_1，ξ_2正态分布时(表 3)，乘积指标法在 N = 100 或 200 时统计检验力低于 0.8，N = 500 时统计检验力接近或高于 0.8；LMS 和无信息贝叶斯法在 N = 500 的 3 种相关条件、和 N = 200 且ϕ_{12} = 0.7 时高于 0.8，其他 5 种条件下低于 0.8；有信息贝叶斯法在所有条件下均接近或高于 0.8。总之，交互效应的检验力是乘积指标法的最低，需要较大的样本才能保证检验力高于 80%；BI 的检验力最高，即使样本容量为 100，也有不错的检验力；LMS 和 BN 的检验力介于上述两种方法之间。

潜变量ξ_1，ξ_2非正态分布时(表 4)，统计检验力的结果与正态情形有类似的结果。所有方法的统计检验力都随着样本容量增加而提高，也随着交互项的非正态性严重程度(ξ_1和ξ_2相关)变大而提高。

5 讨论

当模型包含交互效应时，通常的标准化估计是不恰当的。温忠麟等(2008，也见 Wen et al., 2010)给出了潜变量交互效应恰当标准化估计，并证明了其尺度不变性。本研究系统比较了在外生潜变量不同分布、相关大小和不同样本容量条件下，乘积指标法、LMS和两种贝叶斯法得到的潜变量交互效应恰当标准化估计的表现。

5.1 模拟研究结果

表 5 总结了不同方法得到的潜变量交互效应标准化估计的模拟研究结果，并与原始估计的不同估计方法比较结果(温忠麟, 刘红云, 2020, p202-203)放在一起，方便看出异同。

LMS 和两种贝叶斯法的适当解比例在各种条件下都等于或者接近 100%。乘积指标法的适当解比例随着样本容量和外生潜变量相关的增大而增大，在小样本(N = 100)时适当解比例较低；但在大样本时(N = 500)适当解比例接近 100%。这与前人的结果类似(如 Cham et al., 2012)。

在外生潜变量为正态分布时，乘积指标法和 LMS 得到的交互效应的恰当标准化估计在估计偏差、标准误估计偏差、I类错误率和统计检验力的表现与原始估计的(Cham et al., 2012; Jackman et al., 2011; Kelava et al., 2011; Marsh et al., 2004; 温忠麟, 刘红云, 2020)类似。具体来说，对于乘积指标法，交互效应的估计偏差在大样本时可以忽略，Satorra-Bentler 法矫正的标准误的偏差在大样本时也可以接受，I类错误率在所有情形均可接受，虽然统计检验力比较低，但在大样本时高于 80%。对于 LMS，交互效应的估计偏差在各种条件下均可忽略，标准误的偏差均可接受，且在中等以上样本(N = 200 或以上)时估计偏差和标准误偏差几乎都比其他方法的小。LMS 的I类错误率往往偏高，统计检验力也明显高于乘积指标法。

无先验信息时，贝叶斯法得到的交互效应恰当标准化估计的偏差和标准误偏差在各种条件下都可接受，I类错误率几乎都可接受，统计检验力的表现随样本容量增大和外生潜变量相关增大而变高，在大样本或高相关时都超过 80%。正确设置先验信息时，贝叶斯法得到的交互效应恰当标准化估计的偏差在各种条件下都可接受，且在小样本时比起其他方法的更小；但在各种情形都一致地高估标准误，不过这并非坏事，说明参数估计的实际波动比估计的还小；I类错误率和统计检验力在各种条件下都可接受，且比其他方法的表现要好。

表 5 乘积指标法、LMS 和贝叶斯法的原始估计和恰当标准化估计特点比较

特 点	原始估计				恰当标准化估计			
	无约束的乘积指标法	LMS	贝叶斯法		无约束的乘积指标法	LMS	贝叶斯法	
			无先验	有先验			无先验	有先验

效应估计偏差	正态	*N* 大可接受	可接受	-	可接受	*N* 大可接受	可接受	可接受	可接受
	非正态	*N* 大可接受	偏差较大	-	-	*N* 大可接受	外生潜变量低相关时可接受		
标准误偏差	正态	*N* 大可接受	可接受	-	可接受	*N* 大可接受	可接受	可接受	**高估**
	非正态	*N* 大可接受	偏差可能大	-	-	*N* 大可接受	**可接受**	可接受	*N* 大可接受
I 类错误率	正态	可接受	可接受	-	-	可接受	***N* 大可接受**	可接受	可接受
	非正态	可接受	偏高	-	-	可接受	偏高	样本小或外生潜变量相关低时可接受	
检验力	正态/非正态	*N* 大可接受	较高	-	-	*N* 大可接受	较高	较高	高
标准化估计获得方式						用原始估计和通常标准化估计简单手工计算，或者使用附录 1 的 Mplus 程序	Mplus 编程，输出标准化解 (Mplus 8.2 以上版本)		

注：-表示该条件下未见到有原始估计的模拟研究；粗体表示原始估计和恰当标准化估计不相同的情形。

前人的研究也发现有先验信息的贝叶斯法高估了参数的标准误。Holtmann 等(2016)发现设定正确或错误的先验信息都会高估参数的标准误，且先验信息越强，标准误的偏差越大；但 Lee 等(2007)发现正确设定先验信息比误设先验信息可以更准确地估计标准误，这可能是由于在他们的研究中对先验分布中的方差有不同的设定。

在外生潜变量为非正态分布时，乘积指标法和 LMS 的恰当标准化估计在估计偏差、标准误估计偏差、I类错误率和统计检验力的表现与原始估计的(Cham et al., 2012; Kelava, Nagengast, 2012; Marsh et al., 2004; 温忠麟, 刘红云, 2020)也类似。乘积指标法对交互效应的估计偏差在大样本时可以接受，且比其他方法稳健；标准误的估计偏差在大样本或外生潜变量高相关时可以接受；I类错误率在所有情形均可接受，且在大样本或外生潜变量相关高相关时比其他方法的低；统计检验力在大样本时高于 80%。而 LMS 对交互效应的估计偏差在外生潜变量不相关 ($\phi_{12} = 0$) 时才可以接受，而在高相关估计偏差很大。LMS 的I类错误率几乎都偏高，相应地统计检验力也较高。值得注意的是，LMS 的估计偏差在外生潜变量存在中或高的相关时不受样本容量的系统影响，与 Marsh 等(2004)对另一种分布分析方法 QML 的发现一致。Cham 等(2012)也发现 LMS 在非正态情形稳健性很差。与本文研究有所不同的是，Cham 等(2012)由于使用了峰度和偏度更高的非正态分布， LMS 的表现比本文的还差(更加不稳健)，而乘积指标法在大样本条件下表现依然稳健。

两种贝叶斯法对交互效应的估计偏差在外生潜变量低相关 ($\phi_{12} = 0$) 时才可以接受，小样本且外生潜变量低相关时两种贝叶斯法的估计偏差都较小，而在高相关时的估计偏差很高。无信息贝叶斯法的标准误偏差可以接受，但有信息贝叶斯法在样本大时才可接受。两种贝叶斯法的

I类错误率和统计检验力都随样本或交互项的非正态性严重程度(ξ_1和ξ_2相关越高，非正态性越严重)增高时提高。

总的来说，各种方法得到的潜变量交互效应的标准化估计与原始估计的表现非常类似，说明选择估计方法时不用刻意区分是要计算标准化估计还是原始估计。

5.2 贝叶斯法的表现

由于贝叶斯法结合了先验信息与观测数据，被认为在小样本下具有优势 (Gelman et al., 2014; Lee & Song, 2004; Lee et al., 2007; Muthén & Asparouhov, 2012)。但在潜变量交互效应估计方面，对贝叶斯法的研究相对其他方法要少，尤其是在 I 型错误率和检验力方面，几乎是空白(见表 5)。

本研究发现，贝叶斯法在小样本情形并不一定比 LMS 的估计准确，无先验信息的贝叶斯法在多数情形比 LMS 略差。只有加入了准确的先验信息后，贝叶斯法才表现出一定的优势，但有信息贝叶斯法会高估标准误。这与最近 Smid 等(2020)的结论一致：在无先验信息时，贝叶斯法可能会出现严重的偏差，比极大似然估计更差，只有先验信息准确时，贝叶斯法才表现出小样本情形的优良特性。

当先验信息准确时，先验分布的方差越小，估计结果越准确，检验力越高(也见 Fang et al., 2019; Miočević et al., 2017)。但若设置的先验信息不准确，先验分布的方差越小，估计结果越不准确，只有当样本量较大时，才能弥补错误的先验信息带来的问题(Fang et al., 2019; Lee et al., 2007)。在实际的研究中，可以基于前人的研究或元分析等得到的相关参数设定先验信息。若对这些参数值的把握不大，可以对先验分布设定较小的精度，例如方差大于 1 (Fang et al., 2019; Miočević et al., 2017)。而当较难获得参数的先验信息时，可以首先使用其他方法如乘积指标法、LMS 对潜变量交互效应模型进行估计，再基于得到的估计值来设置先验分布。

其他场合的一些研究发现，贝叶斯法在非正态情形(包括外生潜变量相关)有稳健性(Gelman et al., 2014; Kelava et al., 2008; Kelava & Nagengast, 2012)。但我们的结果表明，在非正态(尤其是外生潜变量高相关)情形下，贝叶斯法没有表现出优势。在外生潜变量正态但高相关时，虽然贝叶斯法比乘积指标法的表现要好，但与 LMS 相比并无优势；而在外生潜变量非正态分布时，贝叶斯法与 LMS 的交互效应估计都有很高的估计偏差，而此时乘积指标法表现得比较稳健。其他研究者在非正态情形估计结构方程模型时，也发现贝叶斯法并不稳健(Jia, 2016; Lee & Song, 2004)。

分析潜变量交互效应时，如何判断非正态性是否严重呢？首先是检查外生潜变量的指标是否有严重非正态。因为潜变量及其指标只相差一个误差，如果指标是正态的，那么相应的潜变

量也是正态的，反之亦然。其次是看外生潜变量之间的相关大小，ϕ_{12}越大，乘积项$\xi_1\xi_2$的非正态性越严重。

5.3 给应用研究者的建议

在变量正态分布时，推荐使用 LMS 估计潜变量交互效应，对交互效应及其标准误的估计在中、大样本时最为准确，而且在小样本或潜变量高相关情形依然表现良好，但要小心第 I 类错误，应当结合效应量做出推断。如果能获得比较准确的先验信息，则可以考虑使用贝叶斯法，尤其是在小样本情形，估计偏差小、检验力高、I 类错误率可以接受，但往往会高估标准误，因而应当避免直接使用标准误的估计。无先验信息时贝叶斯法与 LMS 相比并没有表现出优越性。乘积指标法在大样本时对交互效应及其标准误的估计也可接受，但统计检验力较低，对于检验不显著的结果需要慎重。

在变量非正态分布时，推荐使用无约束的乘积指标法，但样本容量要大，最好不小于500(此时交互效应估计偏差及标准误偏差、统计检验力和 I 类错误率都可接受)。若外生潜变量之间相关低(可以通过验证性因子分析去估计和检验)，小样本时可以考虑使用无信息贝叶斯法，各个评价指标都可接受。

值得注意的是，如果使用 LMS 和贝叶斯法，只要在程序中要求输出标准化解，Mplus8.2或以上版本就会直接输出交互效应的恰当标准化估计；但如果使用无约束的乘积指标法，要么先得到原始解和通常标准化解然后使用公式(5)计算得到，要么使用附录 1 得到（读者不难套用或者修改使用）。

5.4 局限和展望

首先，在我们的研究中，每个因子的指标数目相等，所有指标都具有中等大小的负荷。Marsh 等(2004)的研究结果表明负荷大小对估计方法比较的影响较小。在指标数不等时，Wu 等(2013)研究了乘积指标法与 LMS 得到的潜变量交互效应的估计的表现，与本文结果类似，但对恰当标准化估计，贝叶斯法的表现以及与 LMS 方法的比较还需要进一步研究。

其次，在某些情况下，随样本容量增大贝叶斯估计偏差先增大后下降，这种现象在其他的贝叶斯结构方程模型的估计中也有类似的发现(Holtmann et al., 2016; Lee et al., 2007; van Erp et al., 2018; Yuan & MacKinnon, 2009)。对于有先验信息的情形，这可能是先验信息与样本容量对后验分布的影响相互拮抗的结果：当样本容量很小时，准确的先验信息可以帮助估计后验分布，而样本容量很大时，样本信息的作用更大(Lee et al., 2007)。随着样本容量由小变大，先验信息的作用减弱，而样本容量的信息还不够准确，这时的参数估计可能较小样本和大样本都更差。无先验信息时贝叶斯估计也出现了类似的现象，可能表明贝叶斯法在样本容量不高时估计的稳定性有限，这方面还需要更深入的研究。

近年来，研究者提出了一些新方法应对潜变量交互效应模型面临的不同问题。在混合模型框架下，有研究者使用贝叶斯估计(Kelava & Nagengast, 2012)或极大似然估计(Kelava et al., 2014)分析非线性的潜变量模型，结果表明这些方法在非正态情形下有较小的偏差，但比 LMS 的复杂性高(Asparouhov & Muthén, 2020)，还会面临合理选择潜类别个数的问题(Kelava & Nagengast, 2012)。Brandt 等(2018)提出了贝叶斯 Lasso 方法估计包含多个交互效应和二次效应的模型，Aytürk 等(2020)使用 LMS 分析指标为定序变量的潜变量交互效应模型。这些新方法在建立标准化模型时都应当考虑恰当的标准化模型和标准化估计。

参 考 文 献

Algina, J., & Moulder, B. C. (2001). A note on estimating the Jöreskog-Yang model for latent variable interaction using LISREL 8.3. *Structural Equation Modeling: A Multidisciplinary Journal*, *8*(1), 40–52.

Arminger, G., & Muthén, B. (1998). A Bayesian approach to nonlinear latent variable models using the Gibbs sampler and the Metropolis-Hastings algorithm. *Psychometrika*, *63*(3), 271–300.

Asparouhov, T., & Muthén, B. (2010). *Bayesian analysis of latent variable models using Mplus*. Muthén & Muthén.

Asparouhov, T., & Muthén, B. (2020). Bayesian estimation of single and multilevel models with latent variable interactions. *Structural Equation Modeling: A Multidisciplinary Journal*, *28*(2), 314–328.

Aytürk, E., Cham, H., Jennings, P. A., & Brown, J. L. (2020). Exploring the performance of latent moderated structural equations approach for ordered-categorical items. *Structural Equation Modeling: A Multidisciplinary Journal*, *28*(3), 410–422.

Bandalos, D. L. (2002). The effects of item parceling on goodness-of-fit and parameter estimate bias in structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*(1), 78–102.

Boomsma, A. (1983). *On the robustness of LISREL against small sample size and non-normality* [Unpublished doctorial dissertation]. University of Groningen.

Bradley, J. V. (1978). Robustness? *British Journal of Mathematical and Statistical Psychology*, *31*(2), 144–152.

Brandt, H., Cambria, J., & Kelava, A. (2018). An adaptive Bayesian lasso approach with spike-and-slab priors to identify multiple linear and nonlinear effects in structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(6), 946–960.

Brandt, H., Umbach, N., & Kelava, A. (2015). The standardization of linear and nonlinear effects in direct and indirect applications of structural equation mixture models for normal and nonnormal data. *Frontiers in Psychology*, *6*, 1813.

Büchner, R. D., & Klein, A. G. (2020). A Quasi-likelihood approach to assess model fit in quadratic and interaction SEM. *Multivariate Behavioral Research*, *55*(6), 855–872.

Cham, H., West, S. G., Ma, Y., & Aiken, L. S. (2012). Estimating latent variable interactions with nonnormal observed data: A comparison of four approaches. *Multivariate Behavioral Research, 47*(6), 840–876.

Champoux, J. E., & Peters, W. S. (1987). Form, effect size and power in moderated regression analysis. *Journal of Occupational Psychology*, *60*(3), 243–255.

Fang, J., Wen, Z., & Hau, K. -T. (2019). Mediation effects in 2-1-1 multilevel model: Evaluation of alternative estimation methods. *Structural Equation Modeling: A Multidisciplinary Journal, 26*(4), 591–606.

Foldnes, N., & Hagtvet, K. A. (2014). The choice of product indicators in latent variable interaction models: Post hoc analyses. *Psychological Methods*, *19*(3), 444–457.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014). *Bayesian data analysis* (3rd ed.). Boca Raton, FL: CRC press.

Geman, S., & Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *6*(6), 721–741.

Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, *57*(1), 97–109.

Hau, K. -T., & Marsh, H. W. (2004). The use of item parcels in structural equation modelling: Non-normal data and small sample sizes. *British Journal of Mathematical and Statistical Psychology*, *57*(2), 327–351.

Holtmann, J., Koch, T., Lochner, K., & Eid, M. (2016). A comparison of ML, WLSMV, and Bayesian methods for multilevel structural equation models in small samples: A simulation study. *Multivariate Behavioral Research*, *51*(5), 661–680.

Hoogland, J. J., & Boomsma, A. (1998). Robustness studies in covariance structure modeling: An overview and a meta-analysis. *Sociological Methods & Research*, *26*(3), 329–367.

Ito, T. A., Friedman, N. P., Bartholow, B. D., Correll, J., Loersch, C., Altamirano, L. J., & Miyake, A. (2015). Toward a comprehensive understanding of executive cognitive function in implicit racial bias. *Journal of Personality and Social Psychology*, *108*(2), 187–217.

Jaccard, J., & Wan, C. K. (1995). Measurement error in the analysis of interaction effects between continuous predictors using multiple regression: Multiple indicator and structural equation approaches. *Psychological Bulletin*, *117*(2), 348–357.

Jackman, M. G. A., Leite, W. L., & Cochrane, D. J. (2011). Estimating latent variable interactions with the unconstrained approach: A comparison of methods to form product indicators for large, unequal numbers of items. *Structural Equation Modeling: A Multidisciplinary Journal*, *18*(2), 274–288.

Jia, F. (2016). *Methods for handling missing non-normal data in structural equation modeling* [Unpublished doctorial dissertation]. University of Kansas.

Jöreskog, K. G., & Sörbom, D. (1996). *LISREL 8: User's reference guide*. Chicago, IL: Scientific Software International, Inc.

Kelava, A., Moosbrugger, H., Dimitruk, P., & Schermelleh-Engel, K. (2008). Multicollinearity and missing constraints: A comparison of three approaches for the analysis of latent nonlinear effects. *Methodology*, *4*(2), 51–66.

Kelava, A., & Nagengast, B. (2012). A Bayesian model for the estimation of latent interaction and quadratic effects when latent variables are non-normally distributed. *Multivariate Behavioral Research*, *47*(5), 717–742.

Kelava, A., Nagengast, B., & Brandt, H. (2014). A nonlinear structural equation mixture modeling approach for nonnormally distributed latent predictor variables. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*(3), 468–481.

Kelava, A., Werner, C. S., Schermelleh-Engel, K., Moosbrugger, H., Zapf, D., Ma, Y., … West, S. G. (2011). Advanced nonlinear latent variable modeling: Distribution analytic LMS and QML estimators of interaction and quadratic effects. *Structural Equation Modeling: A Multidisciplinary Journal*, *18*(3), 465–491.

Kenny, D. A., & Judd, C. M. (1984). Estimating the nonlinear and interactive effects of latent variables. *Psychological Bulletin*, *96*(1), 201–210.

Klein, A., & Moosbrugger, H. (2000). Maximum likelihood estimation of latent interaction effects with the LMS method. *Psychometrika*, *65*(4), 457–474.

Klein, A. G., & Muthén, B. O. (2007). Quasi-maximum likelihood estimation of structural equation models with multiple interaction and quadratic effects. *Multivariate Behavioral Research*, *42*(4), 647–673.

Laczniak, R. N., Carlson, L., Walker, D., & Brocato, E. D. (2017). Parental restrictive mediation and children's violent video game play: The effectiveness of the Entertainment Software Rating Board (ESRB) rating system. *Journal of Public Policy & Marketing*, *36*(1), 70–78.

Lee, S. Y., & Song, X. Y. (2004). Evaluation of the Bayesian and maximum likelihood approaches in analyzing structural equation models with small sample sizes. *Multivariate Behavioral Research*, *39*(4), 653–686.

Lee, S. Y., Song, X. Y., & Poon, W. Y. (2004). Comparison of approaches in estimating interaction and quadratic effects of latent variables. *Multivariate Behavioral Research*, *39*(1), 37–67.

Lee, S. Y., Song, X. Y., & Tang, N. S. (2007). Bayesian methods for analyzing structural equation models with covariates, interaction, and quadratic latent variables. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(3), 404–434.

Lin, G. C., Wen, Z., Marsh, H. W., & Lin, H. S. (2010). Structural equation models of latent interactions: Clarification of orthogonalizing and double-mean-centering strategies. *Structural Equation Modeling: A Multidisciplinary Journal*, *17*(3), 374–391.

MacCallum, R. C., & Austin, J. T. (2000). Applications of structural equation modeling in psychological research. *Annual Review of Psychology*, *51*(1), 201–226.

Marsh, H. W., Wen, Z., & Hau, K. -T. (2004). Structural equation models of latent interactions: Evaluation of alternative estimation strategies and indicator construction. *Psychological Methods*, *9*(3), 275–300.

Marsh, H. W., Wen, Z., Nagengast, B., & Hau, K. -T. (2012). Structural equation models of latent interaction. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 436–458). New York: Guilford Press.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, *21*(6), 1087–1092.

Miočević, M., MacKinnon, D. P., & Levy, R. (2017). Power in Bayesian mediation analysis for small sample research. *Structural Equation Modeling: A Multidisciplinary Journal*, *24*(5), 666–683.

Moulder, B. C., & Algina, J. (2002). Comparison of methods for estimating and testing latent variable interactions. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*(1), 1–19.

Muthén, B., & Asparouhov, T. (2012). Bayesian structural equation modeling: A more flexible representation of substantive theory. *Psychological Methods*, *17*(3), 313– 335.

Muthén, B., Kaplan, D., & Hollis, M. (1987). On structural equation modeling with data that are not missing completely at random. *Psychometrika*, *52*(3), 431–462.

Muthén, L. K., & Muthén, B. O. (2019). *Mplus user's guide* (8th ed.). Los Angeles, CA: Muthén & Muthén.

Nagengast, B., Marsh, H. W., Scalas, L. F., Xu, M. K., Hau, K. -T., & Trautwein, U. (2011). Who took the "×" out of expectancy-value theory? A psychological mystery, a substantive-methodological synergy, and a cross-national generalization. *Psychological Science*, *22*(8), 1058–1066.

Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye & C. C. Clogg (Eds.), *Latent variables analysis: Applications for developmental research* (pp. 399–419). Thousand Oaks, CA: Sage.

Schermelleh-Engel, K., Klein, A., & Moosbrugger, H. (1998). Estimating nonlinear effects using a latent moderated structural equations approach. In R. E. Schumacker & G. A. Marcoulides (Eds.), *Interaction and nonlinear effects in structural equation modeling* (pp. 203–238). Mahwah, NJ: Erlbaum.

Smid, S. C., McNeish, D., Miočević, M., & van de Schoot, R. (2020). Bayesian versus frequentist estimation for structural equation models in small sample contexts: A systematic review. *Structural Equation Modeling: A Multidisciplinary Journal*, *27*(1), 131–161.

van Erp, S., Mulder, J., & Oberski, D. L. (2018). Prior sensitivity analysis in default Bayesian structural equation modeling. *Psychological Methods*, *23*(2), 363–388.

Wen, Z., Hau, K. -T., & Marsh, H. W. (2008). Appropriate standardized estimates for moderating effects in structural equation models. *Acta Psychologica Sinica*, *40*(6), 729–736.

[温忠麟, 侯杰泰, Marsh, H. W. (2008). 结构方程模型中调节效应的标准化估计. *心理学报*, *40*(6), 729–736.]

Wen, Z., & Liu, H. (2020). *Analyses of mediating and moderating effects: Methods and applications*. Beijing: Educational Science Publishing House.

[温忠麟, 刘红云. (2020). *中介效应和调节效应：方法及应用*. 教育科学出版社.]

Wen, Z., Marsh, H. W., & Hau, K. -T. (2010). Structural equation models of latent interactions: An appropriate standardized solution and its scale-free properties. *Structural Equation Modeling: A Multidisciplinary Journal*, *17*(1), 1–22.

Wen, Z., & Wu, Y. (2010). Evolution and simplification of the approaches to estimating structural equation models with latent interaction. *Advance in Psychological Science*, *18*(8), 1306–1313.

[温忠麟, 吴艳. (2010). 潜变量交互效应建模方法演变与简化. *心理科学进展*, *18*(8), 1306–1313.]

Wen, Z., Wu, Y., & Hau, K. -T. (2013). Latent interaction in structural equation modeling: Distribution-analytic approaches. *Psychological Exploration*, *33*(5), 409–414.

[温忠麟, 吴艳, 侯杰泰. (2013). 潜变量交互效应结构方程: 分布分析方法. *心理学探新*, *33*(5), 409–414.]

Wu, Y., Wen, Z., Hau, K. -T., & Marsh, H. W. (2011). Appropriate standardized estimates of latent interaction models without the mean structure. *Acta Psychologica Sinica*, *43*(10), 1219–1228.

[吴艳, 温忠麟, 侯杰泰, Marsh, H. W. (2011). 无均值结构的潜变量交互效应模型的标准化估计. *心理学报*, *43*(10), 1219–1228.]

Wu, Y., Wen, Z., & Li, B. (2014). Testing the standardized solutions in structural equation models with interaction effects. *Psychological Exploration*, *34*(3), 260–264.

[吴艳, 温忠麟, 李碧. (2014). 潜变量交互效应模型标准化估计中的检验问题. *心理学探新*, *34*(3), 260–264.]

Wu, Y., Wen, Z., & Lin, G. (2009). Structural equation modeling of latent interactions without using the mean structure. *Acta Psychologica Sinica*, *41*(12), 1252–1259.

[吴艳, 温忠麟, 林冠群. (2009). 潜变量交互效应建模:告别均值结构. *心理学报*, *41*(12), 1252–1259.]

Wu, Y., Wen, Z., Marsh, H. W., & Hau, K. -T. (2013). A comparison of strategies for forming product indicators for unequal numbers of items in structural equation models of latent interactions. *Structural Equation Modeling: A Multidisciplinary Journal*, *20*(4), 551–567.

You, J., Deng, B., Lin, M. P., & Leung, F. (2016). The interactive effects of impulsivity and negative emotions on adolescent nonsuicidal self‐injury: A latent growth curve analysis. *Suicide and Life‐Threatening Behavior*, *46*(3), 266–283.

Yuan, Y., & MacKinnon, D. P. (2009). Bayesian mediation analysis. *Psychological Methods*, *14*(4), 301–322.

Zhang, L., Lu, J., Wei, X., & Pan, J. (2019). Bayesian structural equation modeling and its current researches. *Advance in Psychological Science*, *27*(11), 1812–1825.

[张沥今, 陆嘉琦, 魏夏琰, 潘俊豪. (2019). 贝叶斯结构方程模型及其研究现状. *心理科学进展*, *27*(11), 1812–1825.]

附录 1 无约束的乘积指标法分析潜变量交互效应的 Mplus 程序

```
TITLE: Unconstrained PI Approach
DATA: FILE is ex.dat;
VARIABLE: NAMES ARE y1-y3 x1-x6;
USEVARIABLES ARE x1-x6 y1-y3 x1x4 x2x5 x3x6;
DEFINE: x1x4 = x1*x4; x2x5 = x2*x5; x3x6 = x3*x6;  ! 定义乘积指标
ANALYSIS: ESTIMATOR = MLM; ! MLM 指令用 SB 法矫正标准误
MODEL:
! 指定测量模型
eta by y1 y2 y3;
ksi1 by x1 x2 x3;
ksi2 by x4 x5 x6;
! 潜变量乘积项的测量模型
ksi3 by x1x4 x2x5 x3x6;
! 指定结构模型并设定路径系数的标签
eta on ksi1 (g1)
       ksi2 (g2)
       ksi3 (g3);  ! g1-g3 表示结构方程中 ksi1-ksi3 的系数
! 设定方差、协方差和残差的标签
ksi1 (v1);
ksi2 (v2);
ksi3 (v3);   ! v1-v3 表示 ksi1-ksi3 的方差
ksi1 with ksi2 (c12);   ! ksi1 与 ksi2 的协方差
ksi1 with ksi3 (c13);   ! ksi1 与 ksi3 的协方差
ksi2 with ksi3 (c23);   ! ksi2 与 ksi3 的协方差
eta (v4);  ! eta 残差的方差
MODEL CONSTRAINT:
new(h1-h4);  ! 生成新参数
h4 = (g1**2)*v1+(g2**2)*v2+(g3**2)*v3+2*g1*g2*c12+2*g1*g3*c13+2*g2*g3*c23+v4;
! eta 的方差
h1 = g1*((v1/h4)**(1/2));  ! 主效应(ksi1 系数)的恰当标准化估计
h2 = g2*((v2/h4)**(1/2));  ! 主效应(ksi2 系数)的恰当标准化估计
h3 = g3*((v1*v2/h4)**(1/2));  ! 交互效应(ksi1*ksi2)系数的恰当标准化估计
```

Standardized Estimates for Latent Interaction Effects: Method Comparison and Selection Strategy

WEN Zhonglin, OUYANG Jinying, FANG Junyan

(*School of Psychology & Center for Studies of Psychological Application, South China Normal University, Guangzhou* 510631, *China*)

Abstract

Analyzing the interaction effect of latent variables has become an important topic in both theoretical and empirical studies. Standardized estimation plays an important role in model interpretation and effect comparison. Although Wen et al. (2010) has formulated the appropriate standardized estimation for the latent interaction effects, there is no popular commercial software that provides the appropriate standardized estimation before the launch of M*plus* 8.2 in 2019.

Previous comparisons of methods for estimating latent interaction were based on the original estimation. In this study, through a simulation experiment, the appropriate standardized estimation of latent interaction effects is obtained respectively by four methods: the product indicator (PI) approach, Latent Moderated Structural Equations (LMS), Bayesian method without prior information (BN), and Bayesian method with prior information (BI). Then these estimations are compared in terms of the bias of estimation, the bias of standard error, type I error rate and statistical power.

The true model in the simulation is based on the structural equation $\eta = 0.4\xi_1 + 0.4\xi_2 + \gamma_3\xi_1\xi_2 + \zeta$ where the latent variables η, ξ_1, ξ_2 each had three indicators with a standardized factor loading of 0.7. Experiment factors include the distribution of two exogenous latent variables (normal, non-normal), correlation ϕ_{12} between two exogenous latent variables (0, 0.3 and 0.7), interaction effect γ_3 (0, 0.2), sample size *N* (100, 200, and 500) and estimation method (PI, LMS, BN, BI).

There are five main findings. (1) the proportion of proper solution of LMS and the two Bayesian methods were close to 100% in all treatments, while PI was almost fully proper when *N* = 500. (2) Under the normal condition, the bias of standardized estimation of latent interaction obtained by LMS, BI and BN was ignorable, and PI was acceptable when *N* = 500. Under the non-normal condition, the bias of LMS and Bayesian methods inflated seriously with increasing correlation of two exogenous latent variables, but PI was still acceptable when *N* = 500. (3) Under both distribution conditions, the bias of standard error of standardized estimation of latent interaction obtained by LMS and BN was small and acceptable, while PI was acceptable when *N* = 500, and BI tended to overestimate the standard error. (4) Under normal conditions, the type I error rates of LMS were acceptable only when the sample size was large, while the other methods were acceptable in all conditions. Under the non-normal condition, the

type I error rates of PI were still acceptable, while the other methods were acceptable only when the sample size was small or the correlation between two exogenous latent variables was low. (5) The statistical power of latent interaction obtained by PI was lower than that by any other method, and a large sample size (e.g., N=500) was required to ensure the PI with statistical power over 80%; LMS and BN had higher statistical power, while BI had the highest one in all conditions.

For the latent interaction, the results of comparing different methods in standardized estimation are quite similar to those in the original estimation. Under the normal condition, it is recommended to use LMS to estimate the interaction effect of latent variables, with the caution of Type I error rate and effect size for inference. If accurate prior information can be obtained, Bayesian method is preferred, especially in the case of a small sample. When the variables are not normally distributed, the unconstrained product indicator approach is recommended, which is more robust than the other methods, but the sample size should be large enough (N =500 or above). If the correlation between exogenous latent variables is low (it can be estimated and tested by confirmatory factor analysis), Bayesian method without prior information can be considered for small samples.

Key words latent variable, interaction effect, standardized estimation, product indicator, Latent Moderated Structural Equations, Bayesian estimation